PHILIPS SONY

SYSTEM DESCRIPTION CD-ROM XA

System Description CD-ROM XA

Copyright

For any further explanation of the contents of this document or in case of any perceived inconsistency or ambiguity of interpretation please consult:

Philips Consumer Electronics B.V.
Coordination Office Optical Media Systems
Building SWA-8
P.O. Box 80002
5600 JB Eindhoven
The Netherlands

Tel.: +31 40 736409
Fax.: +31 40 ~~732113~~ Changed into +31 40 2732113

This page is intentionally left blank

Table of Contents

System Description CD-ROM XA

		TAB
Chapter I	Introduction	I
Chapter II	Disc Format	II
Chapter III	Data Retrieval Structure	III
Chapter IV	Audio Data Representations	IV
Chapter V	Video Data Representations	V
Appendix I	CD-ROM XA Disc with Mode 1 Track	A I
Appendix II	Directory Identifiers	A II

This page is intentionally left blank

Table of Contents

		Page
I.1	Scope	1
I.2	Conventions	2
I.3	Recommended vs. Mandatory	4

This page is intentionally left blank

I.1 Introduction

I.1 **Scope**

The "XA" in CD-ROM XA stands for Extended Architecture, i.e. the extension of the basic capabilities of CD-ROM[1] with the addition of a selection of elements (e.g. ADPCM coded audio and some video modes) coming from CD-I[2].

The CD-ROM XA format brings data format compatibility with most host operating system environments to a large variety of multimedia applications.

CD-ROM XA is an extension of the CD-ROM specification, using elements of the CD-I specification, and is in general, consistent with ISO 9660[3].

CD-ROM XA is also a bridge format between CD-ROM and CD-I for publishers, making possible the re-use of CD-ROM XA data in the creation of discs, which can be utilized on a number of selected target systems.

Note: Those items which are not specified in this specification shall follow the CD-DA[4], CD-ROM, CD-I and ISO 9660 specifications.

[1] CD-ROM = Compact Disc Read Only Memory, specified in the System Description of Compact Disc Read Only Memory ("Yellow Book"), N.V. Philips and Sony Corporation.

[2] CD-I = Compact Disc Interactive, specified in the CD-I Full Function Specification ("Green Book"), N.V. Philips and Sony Corporation.

[3] ISO 9660 = Volume and File Structure of CD-ROM for Information Interchange. Ref. No. ISO 9660: 1988 (E).

[4] CD-DA = Compact Disc Digital Audio, specified in the System Description Compact Disc Digital Audio ("Red Book"), N.V. Philips and Sony Corporation.

I.2 **Conventions**

Unless otherwise indicated in this document the conventions used are as follows:

Bit ordering

In this specification, the graphical representation of all multiple-bit quantities is such that the most significant bit is on the left and the least significant bit is on the right.

Figure I.1 Bit ordering for 8 bits (example)

msb lsb

b7	b6	b5	b4	b3	b2	b1	b0

← 1 byte →

Byte ordering

Numerical values represented in binary notation by more than 8 bits (e.g. 16 or 32 bits) are recorded on disc in a field of a descriptor conforming to ISO 9660, chapter 7. The applicable format is specified in the description of the descriptor fields.

I.2 Conventions

Strings
Strings are always given between quotation marks "___".

Hex
All Hexadecimal values are preceeded by a $.

Binary

All binary values are preceeded by a %.

Reserved

All bits or bytes defined as reserved in this document should be set to zero unless otherwise specified.

I.3 **Recommended vs. Mandatory**

All specifications given in this document are **Mandatory** for CD-ROM XA discs unless specifically defined as **Recommended.**

Table of Contents

		Page
II.1 General		1
1.1	Scope	1
1.2	Real Time Sector	1
1.3	Real Time File	1
1.4	Real Time Record	2
1.5	Synchronization	2
II.2	Disc Specification	3
2.1	Track Organization	3
2.2	Lead-in Area	4
2.3	Track Layout in Program Area	5
2.4	Lead-out Area	7
2.5	Subcode P and Q channels	8
II.3	Track Specification	10
3.1	Track Layout	10
II.4	Sector Specification (Mode 2)	11
4.1	Sector Layout	11
4.1.1	General	11
4.1.2	Byte Ordering	11
4.2	Header Field	12
4.3	Subheader Field	13
4.3.1	General Layout	13
4.3.2	Subheader Definitions	13
4.3.2.1	File Number	13
4.3.2.2	Channel Number	14
4.3.2.3	Submode	14
4.3.4.4	Coding Information	17
4.4	General Data Classes	19

This page is intentionally left blank

Table of Contents

		Page
4.5	Form 1 Sector	20
4.5.1	General Layout	20
4.5.2	Error Detection Code and Error Correction Code Fields	21
4.6	Form 2 Sector	22
4.6.1	General Layout	22
4.6.2	Reserved Field	22
4.7	Empty Sector	23
II.5	Encoder Model	24
5.1	Disc Encoder Model	24
5.2	Physical Sector Formatter	25
II.6	Decoder Model	26
6.1	Disc Decoder Model	26
6.2	Sector Processor	27
II.7	Transmission Model for Real Time Files	29

This page is intentionally left blank

List of Illustrations

Figure No.	Caption	Page
II.1	Track Layout	5
II.2	CD-ROM XA Disc with CD-DA Tracks	6
II.3	Example of the general structure and coding of the TOC for a CD-ROM XA disc with one CD-ROM XA track and two CD-DA tracks	9
II.4	Layout of Header Field	12
II.5	Layout of Subheader Field	13
II.6	Bit Encoding for the Subheader Byte	15
II.7	Coding Information Byte	17
II.8	Real Time Bit constraints	19
II.9	Layout of Form 1 Sector	20
II.10	CD-ROM XA Sector Layout	21
II.11	Layout of Form 2 Sector (a)	22
II.12	Layout of Form 2 Sector (b)	22
II.13	Disc Encoder Model	24
II.14	Physical Sector Formatter	25
II.15	Decoder Model	26
II.16	Sector Processor	27
II.17	Transmission Model for Real Time Files	29

This page is intentionally left blank

II.1 **General**

1.1 **Scope**

This chapter gives a specification of the physical CD-ROM XA disc format. Those parts not specified in this chapter are specified in the CD-DA specification or the CD-ROM specification, or in other chapters of the CD-ROM XA specification.

1.2 **Real Time Sectors**

A "Real Time" sector is a new type of CD-ROM sector needed for retrieval of Audio data. The "Real Time" (RT) attribute is given to sectors that by their nature must be retrieved sequentially from a CD-ROM XA disc at its maximum throughput rate, without being interrupted, in order to avoid performance degradation.

As the CD-ROM XA disc has a higher throughput rate than required for Audio due to the use of ADPCM data compression, there is room for other information (i.e. Video, Text and Data) to be recorded in otherwise unused sectors.

For RT sectors, the priority is given to the real time aspect (and not to data integrity). Therefore, only CD-ROM XA systems with built-in support that enables them to correct errors fully transparently in real time will correct errors in these sectors.

1.3 **Real Time File**

A file containing RT sectors is called a Real Time File (RT file). A RT file may be interleaved with other files. Due to the continuous reading of sectors from the disc, the RT files processing is different from traditional data processing. To support the CD-ROM XA controller system to process the RT data correctly without interrupts there is special information in the Subheader. The controller system has to select sectors read from the disc on File and Channel number. It also has to be able to signal to the Operating System and/or the application the occurrence of Trigger, End-of-Record (EOR) and End-of-File (EOF) bits.

1.4 **Real Time Record**

A Real Time Record consists of an integral number of RT sectors of a RT file. The smallest possible RT record is one sector. A RT record is the smallest portion of RT file that can be played.

The playing of a RT record starts at its beginning and usually stops at its end. The last sector of a RT record must have the EOR bit set (see II.4.3.2.3).

RT records may contain several channels interleaved together, (e.g. one video channel with several audio channels in different languages).

1.5 **Synchronization**

The playing back of RT records is an asynchronous operation; the application program executes concurrently while audio playback is in progress.

RT records may contain audio, video and computer data. Audio sectors are sent to the ADPCM processor at a constant rate depending on audio mode, and other information (e.g. video) has to be synchronized to this rate. Audio sectors can be considered a clock upon whose timing the other elements must be synchronized.

Video and data sectors normally contain data information which is tolerant of short access delays of a few sectors. There is an indeterminate delay period between the time a play request is issued and the time the first audio sector is processed. This delay is caused by the disc seek and latency delay. Therefore the application may need special signaling of critical timing events.

A signal is a real time software interrupt generated by the device driver when a pre-defined condition is met. It is possible to have a signal generated at the occurrence of a Trigger, EOR or EOF bits in the Subheader, or at a reading of a specific type of sector.

II.2 Disc Specification

2.1 Track Organization

The CD-ROM XA disc conforms to the CD-ROM specification with the following additional specifications.

A CD-ROM XA track is a data track recorded in Mode 2 Form 1 or Form 2 according to this specification.

It is **recommended** to use only one (1) CD-ROM XA track.

In some CD-ROM XA applications, it may be necessary to use more than one CD-ROM XA track. This may be done providing that all of the CD-ROM XA tracks are grouped together.

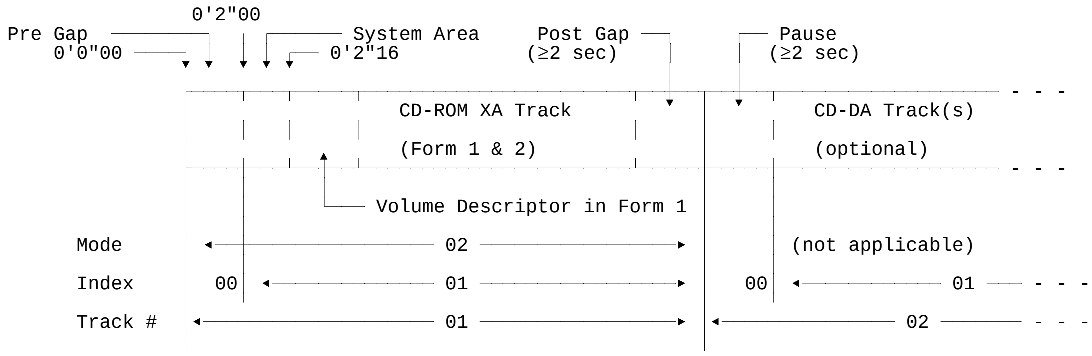
Note: Some early CD-ROM drives are not able to read discs with data tracks recorded in Mode 2. If it is desired that discs be read by CD-ROM drives capable of reading Mode 1 tracks only, a special CD-ROM XA disc can be created by adding a CD-ROM Mode 1 track as a first track. This optional format is described in Appendix I: CD-ROM XA Disc with Mode 1 Track.

2.2 **Lead-in Area**

It is **recommended** that the lead-in area be encoded as a data track (with track number 00) containing empty sectors (see II.4.7) of Mode 2 Form 2 with File Number 0 (zero) and Submode $20.

2.3 **Track layout in Program Area**

Figure II.1 **CD-ROM XA Disc Track Layout**



The CD-ROM XA disc consists of CD-ROM XA Track(s) (Form 1 and Form 2 sectors of Mode 2), followed by CD-DA Track(s) if desired.

The subheader in the Pre and Post Gap has empty sectors with File Number 0 (zero) and Submode 0 (zero).

The Form 1 Post Gap in the last part of the CD-ROM XA track is not required if there is no CD-DA track.

Figure II-2

CD-ROM XA Disc with CD-DA Tracks

Lead-in	Program Area			Lead out
Data	CD-ROM XA	CD-DA	CD-DA	Audio silence
	Track 1	Track 2	Track 3	

In this example the program area contains one CD-ROM XA track and two CD-DA tracks.

2.4 **Lead-out Area**

If the last track in the program area of the disc is a CD-DA audio track, then the lead-out area has to be encoded as a silent audio track.

If the last track in the program area of the disc is a CD-ROM XA track, then the lead-out area has to be encoded as a Mode 2 Form 2 (see II.4.6) data track containing empty sectors (see II.4.7) with File Number 0 (zero) and Submode $20.

2.5 Subcode P and Q channels

The subcode channels have the same technical specifications on a CD-ROM XA disc as is given in the CD-ROM specification with the following exceptions:

(1) The first track on the disc is track number one.

(2) The PSEC of POINT $AO in the lead-in area is used for the identification of the type of disc.

POINT = $AO	PSEC = $00	: means a CD-DA or CD-ROM disc (with first track in Mode 1, see Appendix AI).
	PSEC = $10	: means a CD-I disc
	PSEC = $20	: CD-ROM XA disc (with first track in Mode 2)

All other numbers are reserved.

PFRAME = $00 in all cases.

Figures below give examples of the encoding of the TOC and subcode Q-channel as well as for the general structure of CD-ROM XA discs.

Figure II.3 **Example of the general structure and coding of the TOC for a CD-ROM XA disc with one CD-ROM XA data track and two CD-DA tracks**

		Information Area							
Inner side of the disc		Lead-in Area	Program Area					Lead-out Area	Outer side of the disc
Q CHANNEL	Contr.	Data	Data		Audio		Audio	Audio	
	TNO	00	01		02		03	$AA	
	Index		00	01	00	01	01	01	
	Time		2 sec		≥ 2 sec				
	Point/ ATime	PSEC =$20		*			*	*	
HEADER	Mode	-02	02						
	Mins Header Address	A0							

* These points as specified in the Q-channel give the start position of the first and the last track or lead-out area in ATime with an accuracy of ± one second.

II.3 Track Specification

3.1 Track Layout

A CD-ROM XA track is divided into Mode 2 sectors of 2352 sequential bytes each. There are four different types of CD-ROM XA sectors depending on the type of information they contain i.e. Audio, Video, Data and Empty Sectors. Each sector can be uniquely addressed by a BCD-coded absolute time value in the header field of the sector (see II.4.2).

II.4 Sector Specification

4.1 Sector Layout

4.1.1 General

A CD-ROM XA sector contains the following data fields:

- Sync field	12 bytes
- Header field	4 bytes
- Subheader field	8 bytes
- Data field	2328 bytes

All data in a CD-ROM XA sector, except the synchronization field is scrambled (see CD-ROM specification).

4.1.2 Byte Ordering

The bytes in a CD-ROM XA sector are indexed from 0 to 2351. The index also indicates the sequential order in which the bytes have to be processed by the encoder.

4.2 **Header Field**

The organization of the header field (4 bytes) is given in Figure II.4 below. The header contains the CD-ROM XA sector address (3 bytes) and the mode byte.

The relative time difference of the sector address in the header field to the subcode channel Q ATIME has to be less than one second.

Figure II.4 **Layout of the Header Field**

Byte Number	Byte Value
12	Minutes
13	Seconds
14	Sectors
15	Mode

Minutes = a copy of the data contents of AMIN (subcode channel Q)
Seconds = a copy of the data contents of ASEC (subcode channel Q)
Sectors = a copy of the data contents of AFRAME (subcode channel Q)
Mode = 2, must be used for CD-ROM XA tracks.

4.3 **Subheader Field**

4.3.1 **General Layout**

Subheader bytes consist of 8 bytes: **File Number, Channel Number, Submode and Coding Information** bytes, each double-written for data integrity. (See Figure II.5)

Figure II.5 **Layout of Subheader Field**

Byte Number	Byte Value
16	File Number
17	Channel Number
18	Submode
19	Coding Information
20	File Number
21	Channel Number
22	Submode
23	Coding Information

4.3.2 **Subheader Definitions**

4.3.2.1 **File Number**

The File Number is used to identify all sectors that belong to one file. A given file may be physically interleaved with other files on the disc. All sectors of a logical file have the same value in the File Number byte. The File Number can be used to select sectors that belong to the same logical file and to reject any others.

File Number with a byte value of 0 is only used for a file, directory or any other area on disc (like Path Tables or Volume Descriptors) which are recorded in consecutive sectors on disc. No interleaving is possible when the File Number is 0.

Files with File Number 1 to 255 may be recorded in consecutive sectors or may be interleaved with sectors with different File Numbers.

4.3.2.2 **Channel Number**

An interleaved file may contain several different pieces of information that need be chosen in combination or separately at playback. To facilitate the real-time selection of such information each piece of information may be given a unique Channel Number.

Channel Number 0 to 15 may be used to ADPCM audio sectors, and 0 to 31 for data and video sectors.

4.3.2.3 **Submode**

The Submode byte defines global attributes of a sector as required for the initial selection and allocation of a sector in the system, termination of a file or record, initialization of an additional layer of error correction, and synchronization.

Synchronization:

For synchronization purposes there are three types of marker bits that can be attached to a transmitted sector, namely EOF, EOR and Trigger. These bits can be used by the application software to synchronize the audio, video and data bit streams. The Trigger bit is not affected by the channel allocation (true for all channels) but EOF and EOR are both channel allocated. Therefore it is recommended to use one dedicated channel for EOR (and EOF) markers which have to be used independently of the output channel selection.

Actual handling of some bits, like Real-Time (bit 6) and Trigger (bit 4), may be dependent on the target operating system.

II.4 Sector Specification

The Submode byte is bit-encoded as follows:

Figure II.6 **Bit Encoding for the Submode Byte**

Bit Number	Bit Name
7	End Of File (EOF)
6	Real Time Sector (RT)
5	Form (F)
4	Trigger (T)
3	Data (D)
2	Audio (A)
1	Video (V)
0	End Of Record (EOR)

End Of File (EOF) : The last sector of a file is indicated by bit 7 equal to 1. All other sectors have bit 7 equal to zero.

Real Time Sector (RT) : If bit 6 has the value 1 then the data has to be processed without interrupting the real-time behavior of the CD-ROM XA system. For example, audio sectors have to be transferred to the ADPCM decoder in real time in order to avoid the overflow or underflow of data.

Form (F) : This bit has a value of 0 for all sectors recorded in Form 1 and a value of 1 for all sectors recorded in Form 2.

Trigger (T) : This bit is used to synchronize the application with various coding information, like visuals to audio, in real time. The bit when set to one can be used to generate an interrupt.

Data (D) : This bit has a value of 1 for program related data sectors. Otherwise it is zero. When this bit is set to one, the Form bit must be zero.

Audio (A) : This bit has a value of one for audio sectors. Otherwise it is zero. When this bit is set to one the Form bit is also set to one.

Video (V) : This bit has value of one for video sectors. Otherwise it is zero.

End Of Record (EOR) : This bit must have the value one for the last sector of a logical record. Otherwise it is zero. The use of the EOR bit is only mandatory for real time records.

Only one of bit 1, bit 2 or bit 3 may have the value one at the same time. One of bit 1, bit 2 and bit 3 must be set to the value one for all sectors except empty and message sectors. If bits 1, 2 and 3 of the submode byte are zero then the sectors are empty.

4.3.2.4 **Coding Information**

The Coding Information byte defines the details of the type of data located in the user area of the sector. This depends on whether the data is audio, video or program related data.

Figure II.7 **Coding Information Byte**

<table>
<tr><td>bit
sector</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td><td></td></tr>
<tr><td rowspan="2">Audio</td><td rowspan="2">0</td><td>Emph</td><td colspan="4">ADPCM Level[1]</td><td colspan="2">Mode[1]</td><td rowspan="2"></td></tr>
<tr><td>0 = Off
1 = On</td><td colspan="4">0000 = Level B
0001 = Level C</td><td colspan="2">00 = Mono
01 = Stereo</td></tr>
<tr><td rowspan="5">Video</td><td>0</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td><td>CDI</td></tr>
<tr><td rowspan="4">1</td><td rowspan="3">0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>ASM</td></tr>
<tr><td colspan="3">Resolution[1]</td><td colspan="3">Coding[1]</td><td rowspan="3">EVM</td></tr>
<tr><td colspan="3">000=320x200
001=640x480</td><td colspan="3">001=CLUT1
010=CLUT2
011=CLUT4
100=CLUT8</td></tr>
<tr><td>1</td><td colspan="6">Reserved</td></tr>
<tr><td>Data</td><td colspan="8">Reserved</td><td></td></tr>
</table>

Emph	:	Emphasis bit
Level B	:	4 bit 37.8 KHz sampling rate
Level C	:	4 bit 18.9 KHz sampling rate
CDI	:	CD-I Video Mode
ASM	:	Application Specific Video Mode
EVM	:	Extended Video Mode

[1] Other combinations reserved.

Three on-disc video coding modes are available:

- CD-I Video Mode
  When this mode is applied, it should fully conform to the CD-I Specification.[2]

- ASM or Application Specific Video Mode
  The coding is application specific, i.e. to be interpreted by application software.

- EVM or Extended Video Mode
  This mode is further described in Chapter V.

Note: It is not recommended to use the ASM mode except for special applications. Use of the Extended Video Mode is preferred.

[2] CD-I Full Functional Specification, chapter V.

4.4 **General Data Classes**

Sectors in a CD-ROM XA track may contain two generic classes of data:

. data with a strong requirement for data integrity where concealment is not possible (e.g. text, computer data)
. data where errors can be concealed by some interpolation scheme (e.g. audio, video).

Text and computer data, for example, have a strong requirement for an extra layer of error correction beyond the CIRC (Cross-Interleaved Reed-Solomon Code) used on all Compact Discs. For this class of data, error detection and correction should be implemented according to the CD-ROM XA EDC/ECC. The sector structure used for this class of data is called Form 1 and is specified in detail in II.4.5.

The latter class of data has a strong requirement for maximum data throughput per time unit. The sector structure used for this class of data is called Form 2 and is specified in detail in II.4.6.

It should be noted that both Form 1 and Form 2 may be used for real-time data. However, for Form 1 real time sectors, priority is given to the real time aspect (and not to data integrity). Therefore only CD-ROM XA-systems with built-in hardware support which enables them to correct errors fully transparently in real-time without requiring CPU intervention will correct these errors. Consequently, a CD-ROM XA system without real time ECC handling will not correct Form 1 real time sectors.

Allowed combinations of different sector types and classes are given in the table below.

Figure II.8 **Real Time Bit Constraints**

	Form 1		Form 2	
	RT = 0	RT = 1	RT = 0	RT = 1
Audio	No	No	Yes	Yes
Video	Yes	No	Yes	Yes
Data	Yes	Yes	No	No
Empty	Yes	Yes	Yes	Yes

4.5 **Form 1 Sector**

Form 1 can be used for Data sectors, Video sectors and Empty sectors.

4.5.1 **General Layout**

A Form 1 sector consists of 2352 bytes. Each byte in the sector can be identified by B_i where i = 0 to 2351.

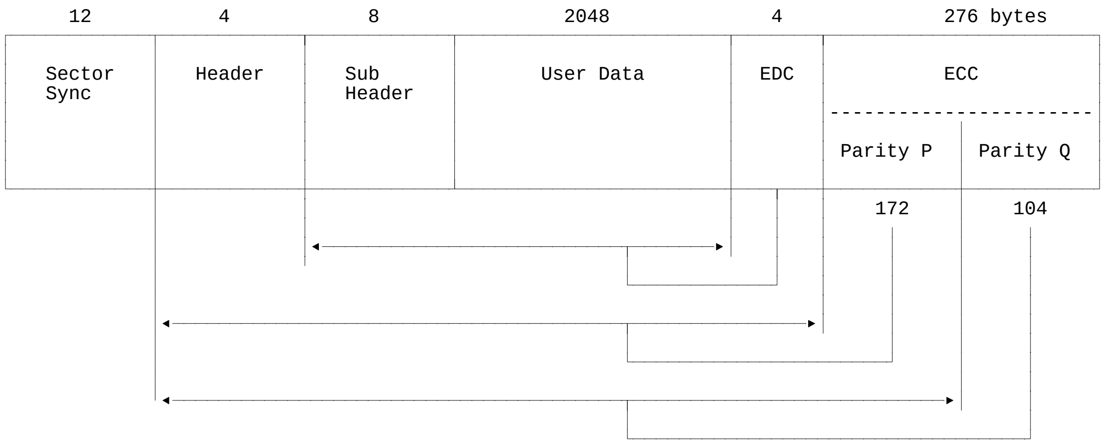
The layout of a Form 1 sector is as follows:

Figure II.9 **Layout of Form 1 Sector**

Byte	Meaning
B_o to B_{15}	The Synchronization and Header fields.
B_{16} to B_{23}	The Subheader field.
B_{24} to B_{2071}	2048 Bytes of data whose content depends on the sector type.
B_{2072} to B_{2075}	EDC field of 4 bytes.
B_{2076} to B_{2351}	ECC field.

4.5.2 **Error Detection Code and Error Correction Code Fields**

Figure II.10 **CD-ROM XA Sector Layout**



Except for the following, EDC and ECC are the same as the Yellow Book:

- The EDC used is a 32 bit CRC on the data field from the beginning of Subheader to the end of the 2048 byte User Data area.

- Parity P of the ECC shall be calculated from the beginning of the Header to the end of the EDC field as the contents of the Header are all zero.

- Parity Q of the ECC shall be calculated from the beginning of the Header to the end of Parity P field as the contents of the Header are all zero.

4.6 **Form 2 Sector**

Form 2 can be used for ADPCM Audio sectors, Video sectors and Empty sectors.

4.6.1 **General Layout**

A Form 2 sector consists of 2352 bytes. Each byte in the sector can be identified by B_i where i = 0 to 2351.

The layout of a Form 2 sector is given in Figure II.11 and II.12.

Figure II.11 **Layout of Form 2 Sector (a)**

Byte	Meaning
B_o to B_{15}	The Synchronization and Header fields.
B_{16} to B_{23}	The Subheader field.
B_{24} to B_{2347}	2324 Bytes of data whose content depends on the sector type.
B_{2348} to B_{2351}	4 Bytes reserved for quality control during the CD-ROM XA disc production process.

Figure II.12 **Layout of Form 2 Sector (b)**

12	4	8	2324	4 bytes
Sector Sync	Header	Sub Header	User Data (ADPCM Audio or Video)	Reserved (0 or EDC)

4.6.2 **Reserved Field**

The reserved field (B_{2348} to B_{2351}) contains 4 bytes that are reserved for quality control during the CD-ROM XA disc production process. These bytes are ignored by the CD-ROM XA system. It is **recommended** that the same EDC algorithm should be used here as is used for Form 1 sectors. If this algorithm is not used, then the reserved bytes are set to zero.

4.7 Empty Sector

An Empty sector may be Form 1 or Form 2 and does not contain any CD-ROM XA data. Empty sectors may also be used in the Lead-in and/or Lead-out Area. In the Program Area Empty sectors can be used to fill up file space particularly for Real-time Files.

The Subheader of an Empty Sector has the following restrictions:

(1) Channel Number must be zero.

(2) The Video, Audio and Data bits in the Submode byte must be zero.

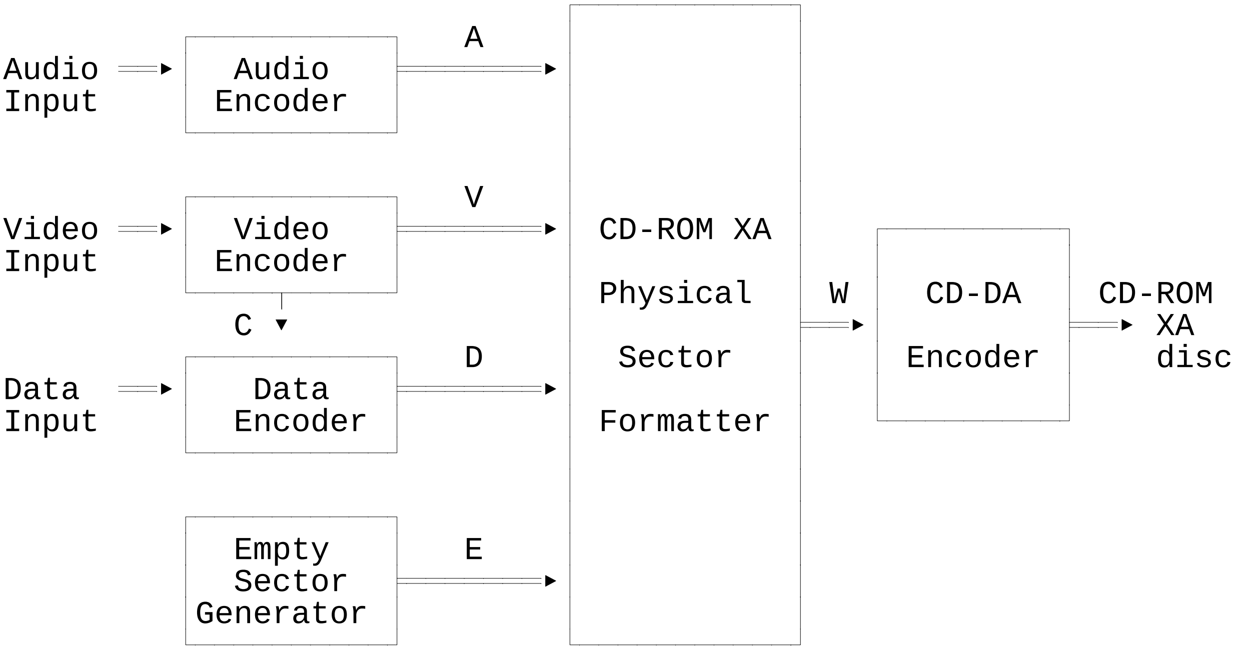
(3) The Coding Information byte must be zero.

It is **recommended** that Empty Sectors are Form 2 and that the User Data field bytes are zero.

II.5 Encoder Model

5.1 Disc Encoder Model

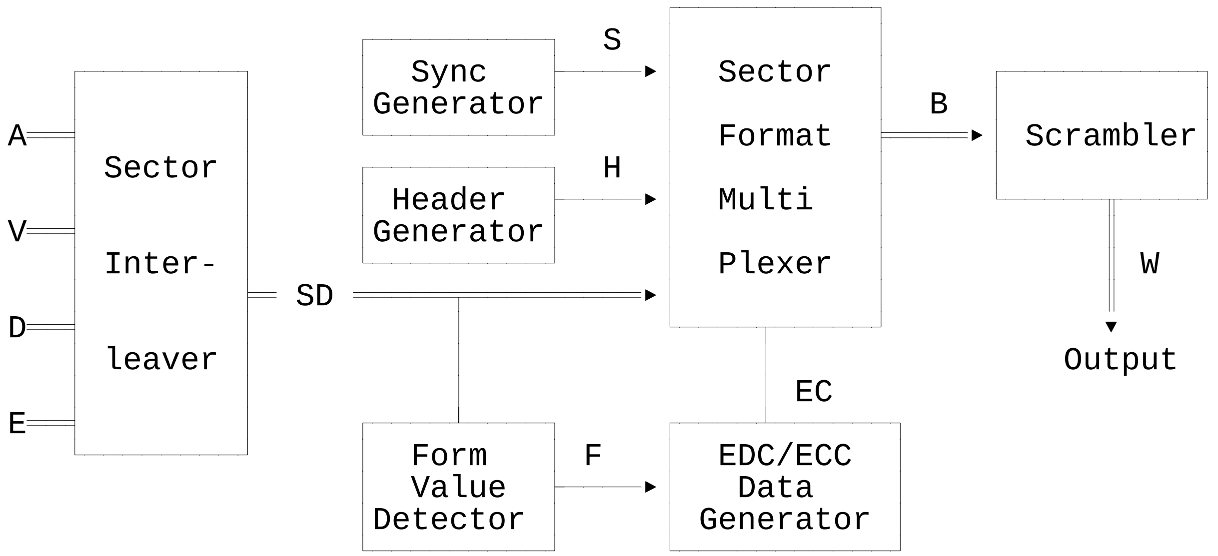
Figure II.13 **Encoder Model**



A = Audio sector
V = Video sector
D = Data sector
E = Empty sector
C = Video control data
W = CD-ROM XA Sector

5.2 **Physical Sector Formatter**

Figure II.14 **Physical Sector Formatter**

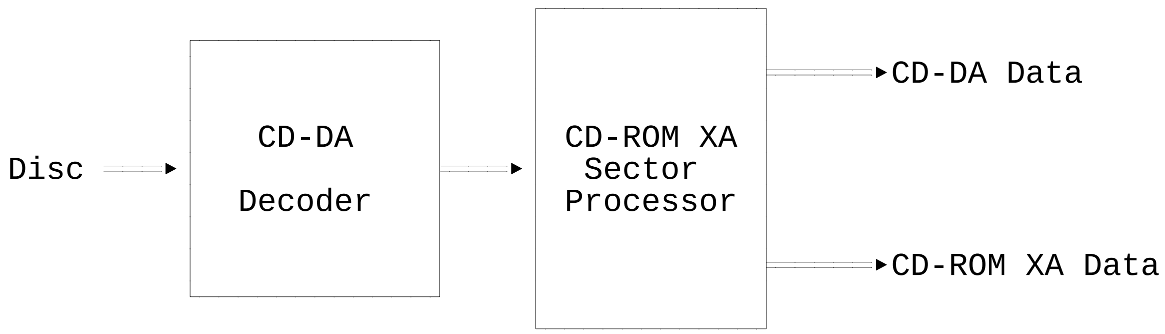


A = Audio sector data
V = Video sector data
D = Data sector data
E = Empty sector data
S = Synchronisation data
H = Header data
SD = Sector data
EC = Error Detection and Correction data
F = Form value
B = Serial bit stream of CD-ROM XA sector data
W = 16-bit word stream of CD-ROM XA sector data to CD-DA encoder

II.6 **Decoder Model**

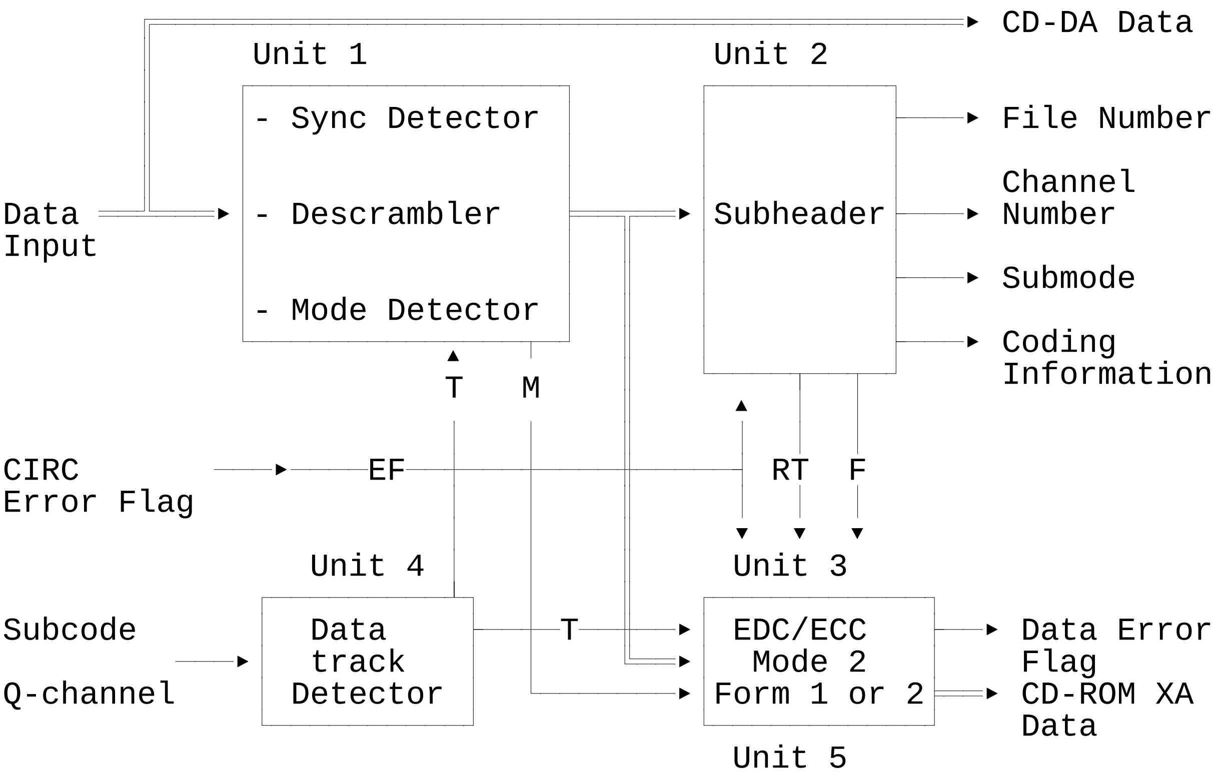
6.1 **Disc Decoder Model**

Figure II.15 **Decoder Model**



6.2 **Sector Processor**

Figure II.16 **Sector Processor**



M	=	Mode 2
RT	=	Real-time Flag of Submode Byte
F	=	Form (1 or 2)
EF	=	Set if data is unreliable
T	=	Set if data track

For CIRC see CD-DA specification pp 27 to 37.

For Subcode Q channel see CD-DA specification page 41.

Unit 1

- The synchronization detector is used for all timing needed in the sector processor.

- The descrambler processes the input data stream into sector data bytes (B_{12} to B_{2351}).

- The Mode detector gives the mode value for the data sector from the Header field.

Unit 2

In Unit 2 the error flag from CIRC is used to select the File Number, Channel Number, Submode and Coding Information bytes from the Subheader field (B_{16} to B_{23}) that are reliable. The Real-time flag and the Form value in the Submode byte are passed to Unit 3.

Unit 3

For Form 1 sectors the EDC and ECC data field is used to correct any erroneous data in the data field. In the case of a Form 2 sector no processing is required. A data error flag is set to one if error correction was not possible for a Form 1 sector or if an EDC error was detected.

Unit 4:

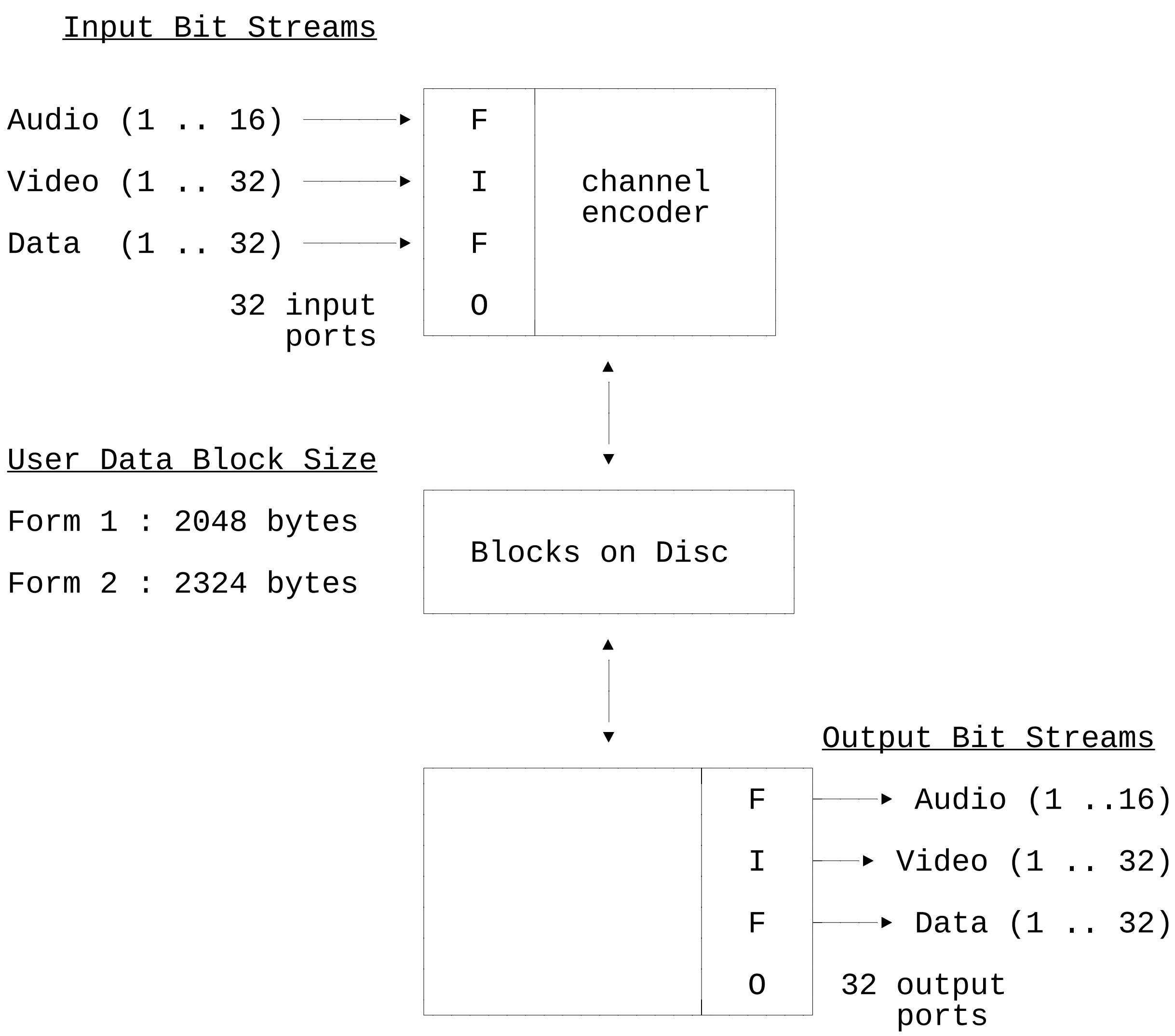
Unit 4 has to detect if the disc is a CD-ROM XA disc or not. This can be done by using the information in the TOC (Subcode Q-channel). It also detects if track is a data track.

Note: The Unit 2 and Unit 3 processing are only valid for CD-ROM XA data tracks.

II.7 **Transmission Model for Real Time Files**

This section clarifies how a CD-ROM XA Real Time File can handle a set of up to 32 parallel bit streams. The figure below shows the transmission model used.

Figure II.17 **Transmission Model for Real Time Files**



Max. Total Throughput Capacity:

The total channel capacity is between 1.23 and 1.39 Mbit/sec depending on mix of Form 1 and Form 2 sectors.

If the total input bit stream capacity is less than these values then Empty Sectors will be inserted by the channel decoder in the channel bit stream on the disc.

Input Bit Streams:

The channel encoder has up to 32 parallel input ports for a mix of audio, video and data input bit streams. Each bit stream has its own channel number (#0 .. #31) and is stored in a block structured FIFO[3].

Restrictions are that each channel can only handle one sort of bit stream (Audio, Video or Data) and that only up to 16 parallel Audio bit streams are allowed.

The sum of all input bit streams may not be more than the total channel capacity on the disc.

The output from the encoder is 75 sectors of Form 1 or Form 2 per second. The channel encoder transmits the user data blocks from the FIFO in the order they become full. If no block is full then an empty block is transmitted. With each sector on the disc the information related to the sector type (Audio, Video, Data or Empty sector), Form and Channel Number are attached.

Output Bit Streams:

The channel decoder receives 75 sectors per second and puts the decoded user data blocks to the output FIFO as the bit streams are addressed by the Data type and Channel Number. Empty sectors are ignored.

Clock and Timing:

The basic clock is 75 Hz, namely the sector rate on the disc. There is also a sample clock of 588 x 75 = 44100 Hz on the disc. The input and output bit streams may have any clock rate, but within the capacity constraints. To avoid problems with FIFO overflow it is recommended to use clocks synchronized with 75 Hz.

[3] FIFO = First In First Out buffer

Table of Contents

		Page
III.1	General	1
III.2	Data Structure	2
2.1	Disc Label	2
2.2	Sector Subheader Format of the Volume Descriptor Sequence	3
2.3	Sector Subheader Format of the Path Table and Directory Sectors	4
2.4	Location of Data Track and CD-Digital Audio Tracks	5
2.5	File Structure Compatibility Considerations	6
2.6	File Description	7
2.7	System Use Information	8
2.8	Startup Directory	11

This page is intentionally left blank

III.1 **General**

A CD-ROM XA file is a file recorded in Mode 2 Form 1 or Form 2 sectors.

A CD-DA file is a file recorded in CD-DA tracks.

The CD-ROM XA file structure is based upon the file structure for CD-ROM discs in Mode 1 specified in ISO 9660 with extensions listed in this chapter. These extensions consist of:

1. A CD-ROM XA disc label located in the application use field of the Primary Volume Descriptor.

2. Additional system use information in the directory records of CD-ROM XA files.

3. Rules outlining what information should be recorded in the subheader when certain file structure records are recorded in Mode 2.

Note: Mode 2 Form 1 sectors are logically equivalent to Mode 1 sectors

III.2 **Data Structure**

2.1 **Disc Label**

CD-ROM XA discs require the presence of an identical CD-ROM XA disc label in all copies of the Primary Volume Descriptor. The label is located at offset 1024 of the Primary Volume Descriptor (byte position 1025) and the format of the label is:

CD-ROM XA LABEL

BP	Field Name	Content
1025 - 1032	Identifying Signature	'CD-XA001'
1033 - 1034	CD-ROM XA Flags	
1035 - 1042	Startup Directory	d-characters[1] or all 0's
1043 - 1050	Reserved	0's

The bits in the CD-ROM XA flags are defined as follows:

Bit		Value Description
Bits 0-15	0	Reserved

The 8 bytes for the Startup Directory name specify in d-characters the name of the directory in the root directory containing startup programs for multiple operating systems. Any unused bytes will be recorded as spaces ($20). If a Startup Directory is not specified, then all bytes of this field shall contain zero.

An example of the CD-ROM XA label:

```
CDXA    DB  'CD-XA001'   ;    8 byte identifying signature
        DW  ?            ;    CD-ROM XA flags
        DB  'STARTUP'    ;    Startup Directory Name
        DB  dup 8 (0)    ;    8 bytes reserved
```

[1] Character coding conforming to ISO 646

2.2 **Sector Subheader Format of the Volume Descriptor Sequence**

All sectors in the Volume Descriptor Sequence are in Mode 2 Form 1. These sectors must have the Subheader format given below:

1. The File Number byte is zero

2. The Channel Number bits are set to zero

3. The EOR and data bits in the Submode byte are set to one and the rest are zero. Additionally the EOF bit in the Submode byte of the last sector of the Volume Descriptor Sequence is set to one.

4. The Coding Information byte is set to zero.

2.3 Sector Subheader Format of the Path Table and Directory Sectors

All sectors in the Path Table and Directory files are in Mode 2 Form 1. These sectors must have the Subheader format given below:

1. The File Number byte is zero

2. The Channel Number bits are set to zero

3. The data bit in the Submode byte is set to one and the rest are zero. Additionally, the EOR and EOF bits in the Submode byte of the last sector of the Directory file or Path Table are set to one.

4. The Coding Information byte is zero.

2.4 **Location of Data Track and CD-Digital Audio Tracks**

Non-CD-DA data (i.e. ADPCM sound, video data, text data, computer programs and binary data) are contained in sectors after the Volume Descriptor Sequence before any CD-DA tracks. The Volume Descriptor Sequence must always be in track 1.

It should be noted that accessing of CD-DA data can only be reliably accessed within plus/minus 1 second, while data can be accessed down to the sector required. An attribute bit in the directory record of a file indicates if the file is an audio track recorded in CD-DA format, and that it can be played through a CD-DA decoder.

2.5 **File Structure Compatibility Considerations**

Not all receiving systems do or can fully conform to the ISO 9660 specification. Certain restrictions based on the receiving system may have to be taken into account to achieve data interchange. If full data interchange is desired the following general guidelines should be adhered to when authoring a CD-ROM XA disc to account for receiving system limitations:

Primary Volume Descriptor:

- The value of the Volume Set Size field (BP 121 - 124) shall be one
- Volume Identifiers should be unique, if at all possible, to aid receiving systems in identifying discs
- Kanji Supplementary Volume Descriptors using shift-JIS should be identified with Volume Flags bit 0 set to one and an escape sequence of "$+:" in the coded character set for descriptor identifiers field.

Directory Record:

- The file flags (BP 26) Record bit, Associated File bit, Protection bit and Multi-Extent bit should be zero
- The File Name should not contain more than 8 d-characters
- The File Name Extensions should not contain more than 3 d-characters
- There should not be more than one version of a file.

Path Table Record:

- The Directory Identifier should not contain more than 8 d-characters.

Note: Guidelines for specific target receiving systems will be issued separately in the form of Application Notes.

2.6 **CD-ROM XA File Description**

Individual sectors of an interleaved CD-ROM XA file may contain audio, video or ordinary computer data, but these different types of data cannot be mixed within a single sector. The ordering of the audio, video or data sectors is not restricted and is determined according to the requirements of the application. When an interleaved file is read, sectors are read and directly routed to the appropriate hardware (as in the case of audio data) or routed to the application.

The sectors between consecutive audio sectors can be used in a variety of ways. Video (i.e. stills) or text data does not necessarily have such critical timing and can be placed within the gaps. Also, a number of mutually exclusive audio channels can be interleaved within a file to provide different sound tracks, for example, each in a different language.

Directory files must not be interleaved.

2.7 **System Use Information**

All CD-ROM XA Mode 2 and CD-DA files must include additional system use information in the directory record.

The System Use Extension Information is as follows:

Byte Position	**Field Name**	**Content**	
BP 1-4	Owner ID		
BP 5-6	Attributes		
BP 7	Signature byte 1	$58	'X'
BP 8	Signature byte 2	$41	'A'
BP 9	File Number		
BP 10-14	Reserved	0	

The Owner ID field contains the user identification number of the creator of this file. The owner's group ID is contained in BP 1-2 and the owner's user ID is contained in BP 3-4. Note that both these fields are recorded with the most significant byte first. If an Owner ID is not recorded, then the Owner ID field will be set to zero as specified in ISO 9660.

The attributes are defined as follows:

Bit Number	**Contents**
0	Owner read
1	Reserved
2	Owner execute
3	Reserved
4	Group read
5	Reserved
6	Group execute
7	Reserved
8	World read
9	Reserved
10	World execute
11	File contains Form 1 sectors if set to one
12	File contains Form 2 sectors if set to one
13	File contains interleaved sectors
14	CD-DA file if set to one
15	Directory file if set to one

Table of bit combinations

Bit Number					File Number	Contents
11	12	13	14	15		
1	0	0	0	0	0≤ n ≤255	Mode 2 Form 1 file
1	0	1	0	0	0≤ n ≤255	Mode 2 Form 1 with interleaved sectors
0	1	0	0	0	0≤ n ≤255	Mode 2 Form 2 file
0	1	1	0	0	0≤ n ≤255	Mode 2 Form 2 file with interleaved sectors
1	1	0	0	0	0≤ n ≤255	Mode 2 Form 1, Form 2 file
1	1	1	0	0	0≤ n ≤255	Mode 2 Form 1, Form 2 file with interleaved sectors
1	0	0	0	1	n=0	Mode 2 Form 1 Directory file
0	0	0	1	0	n=0	CD-DA file

If bit 11 is set and bit 12 is not set, then the file is recorded only in Mode 2 Form 1. If bit 14 (CD-DA file) is one, then the File Number field shall be zero. The Directory file bit should agree with the Directory file bit setting in the file flags of the Directory Record. Bits 0-7 are recorded in BP 6 and bits 8-15 are recorded in BP 5.

In case of a recorded Extended Attribute Record, the owner ID and the permission in the attributes in the System Use Extension should agree with the information in the Extended Attribute Record.

The File Number is used to identify all sectors that belong to one file. A given file may be physically interleaved with other files on the disc. All sectors of a file have the same value in the File Number byte. The File Number with a byte value of zero is only used for a file that will be read consecutively. No interleaving is possible with File Number 0. File Numbers 1 to 255 may be read consecutively or may be interleaved with other files.

CD-ROM XA files must begin on sector boundaries. As in the CD-I Specification, real-time interleaved file and CD-DA file length is the number of sectors between the first sector of the file and the last sector of the file (marked with the EOF bit) times a sector length of 2048.

File interleaving as defined in ISO 9660 allows only fixed File Units and fixed Interleave Gaps per Extent. This is specified by BP 27 (File Unit Size) and BP 28 (Interleave Gap Size) of the Directory Record for that Extent.

In principal, CD-ROM XA goes a step further in that files may be interleaved with both varying File Units and Interleave Gaps. This is realized by the File Number concept (see 4.3). To ensure compatibility with ISO 9660 for CD-ROM XA, interleaving is restricted to fixed File Units and fixed Interleave Gaps per Extent.

BP 27 and BP 28 of the Directory Record corresponding to an interleaved file shall be set according to ISO 9660, 9.1.7 and 9.1.8.

2.8 Startup Directory

The CD-ROM XA Disc Label includes a provision for identifying a startup directory. If the disc is targeted for multiple operating environments, the startup directory provides a mechanism for identifying a startup program that will operate in the receiving system.

The Startup Directory shall contain a number of subdirectories having Directory Identifiers which correspond to Generic Operating Environments. These Directory Identifiers are called GOE-Directories.

Each GOE-Directory contains per Specific Operating Environment a subdirectory called SOE-Directory. Typically the subdirectory identifiers correspond to the names of the underlying Application Programming/Application Binary Interfaces (API/ABI) or a release level which has an implication for the API/ABI, e.g. VER11 for OS/2 version 1.1.

The contents of a SOE-Directory is a single Startup File, which is an executable application in the receiving system. By means of that Startup File, the disc may be introduced or presented to the user in an implementation dependent way, e.g. a Startup File may function as an:

- Introduction to the usage and contents of the disc
- Application program to perform the task of data retrieval and presentation
- Installation of additional software components, which upgrade a resident software system for CD-ROM to CD-ROM XA.

An example of a Start-up Directory is given in Figure III.1.

Note: Registered Directory Identifiers are listed in Appendix II.

III.2 Data Structure

Figure III.1 **Example of a Startup Directory**

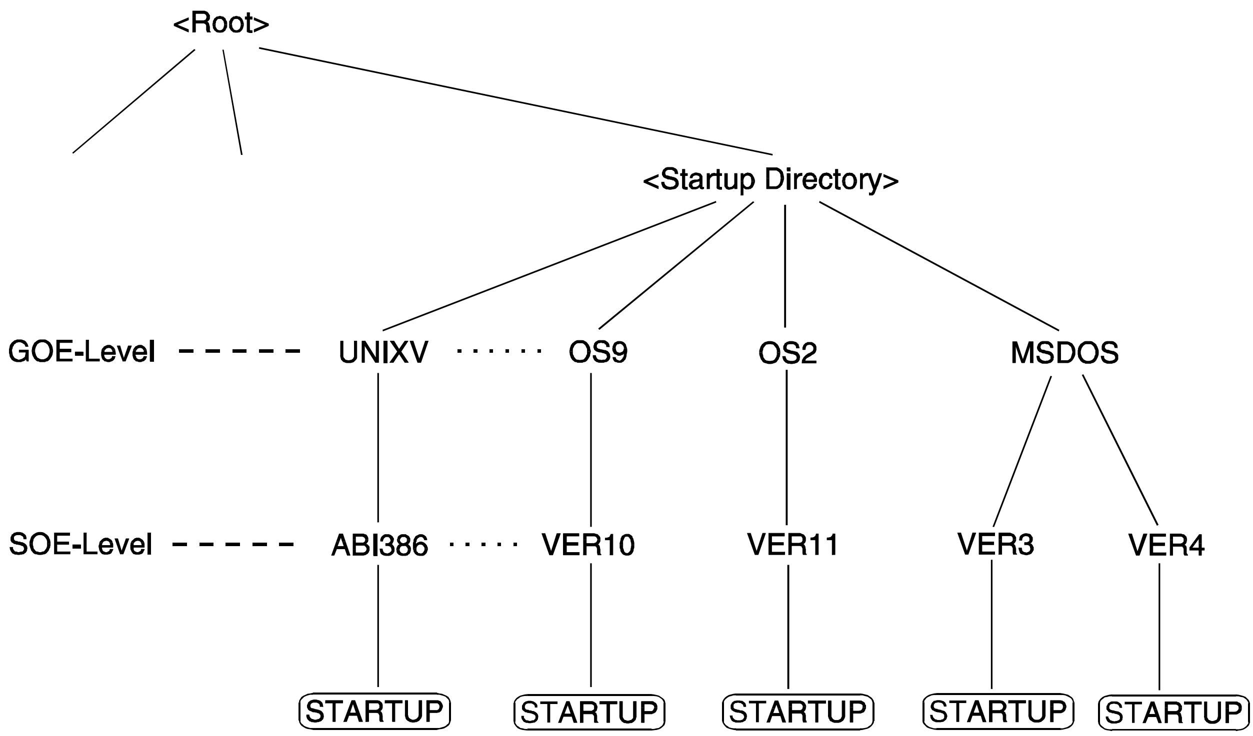


Table of Contents

		Page
IV.1	General Audio Encoding	1
IV.2	Sound Quality Levels	3
IV.3	Audio Sector Data Format	4
	3.1 General	4
	3.2 Subheader Bytes	5
	3.2.1 File Number	5
	3.2.2 Channel Number	5
	3.2.3 Submode	5
	3.2.4 Coding Information	6
	3.3 Audio Block	7
	3.4 Sound Group	8
	3.5 Level B and Level C Audio	9
	3.6 Audio Sector Interleaving	11
IV.4	Audio Data Encoding	12
	4.1 General	12
	4.2 ADPCM Encoder	13
	4.3 Sound Parameters	17
IV.5	ADPCM Decoder	18

This page is intentionally left blank

List of Illustrations

Figure No.	Caption	Page
IV.1	ADPCM Encoding	2
IV.2	CD-ROM XA sound quality levels	3
IV.3	Audio Sector data field	4
IV.4	Encoding for the Submode byte of an Audio sector	5
IV.5	Audio Coding Information byte layout	6
IV.6	Ordering of Sound Groups	7
IV.7	Sound Group layout	8
IV.8	Sound Parameter byte layout	8
IV.9	Sound Parameter layout	10
IV.10	Sound Data layout	10
IV.11	Audio Sector Interleaving	11
IV.12	ADPCM Encoder block diagram	13
IV.13	Predictor Unit	15
IV.14	Noise Shaper Unit	16
IV.15	Gain Values for Sound Parameter Filters	17
IV.16	ADPCM decoder	18

This page is intentionally left blank

IV.1 **General Audio Encoding**

This chapter defines the format and encoding of ADPCM audio data for a CD-ROM XA disc.

Digitally recorded sound in mono or stereo can be encoded as shown in Figure IV.1. The required sound quality level to be encoded has to be set for the sample rate converter and the ADPCM encoder, giving the sampling frequency and number of bits per sample on the disc. The ADPCM encoded audio data is then formatted into an audio sector.

This chapter also defines how the encoded data from a CD-ROM XA disc is decoded back to audible sound.

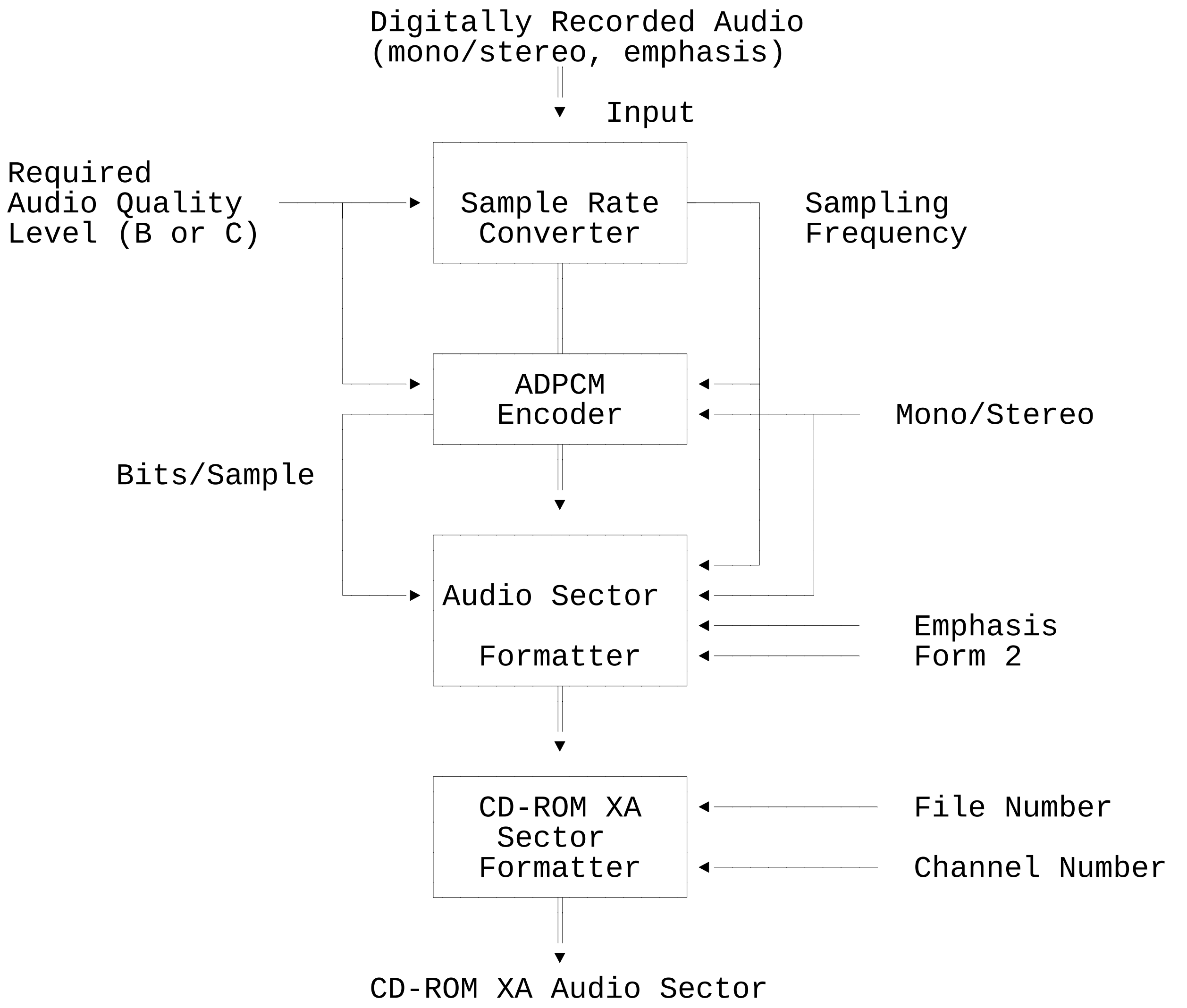
Abbreviations used in this chapter have the following meanings:

ADPCM : Adaptive Delta Pulse Code Modulation
PCM : Pulse Code Modulation
CD-DA : Compact Disc - Digital Audio
S/N : Signal to Noise ratio

Other abbreviations used are defined in the text.

IV.1 General Audio Encoding

Figure IV.1 **ADPCM Encoding**



IV.2 **Sound Quality Levels**

Audio data from more than one audio source can be encoded in a CD-ROM XA file. Each audio source will be identified with an audio Channel Number in the Subheader (see Figure IV.3) of the CD-ROM XA sector.

The encoded sound quality level is dependent on the sampling frequency (fs). The bit rate coming from the CD-ROM XA player is constant, but the maximum available Channel Numbers (N) can vary (see Figure IV.2 below). From the basic audio level parameters, one can derive the maximum:

- audio bandwidth (BW) and
- Channel Numbers (N).

Figure IV.2 gives an overview of the sound quality levels

Figure IV.2 **CD-ROM XA sound quality levels**

Level	fs	b	BW	N	Stereo/ Mono
CD-DA	44.1 KHZ	16	20 kHz	1	Stereo
ADPCM					
Level B	37.8 kHz	4	17 kHz	4 8	Stereo Mono
Level C	18.9 kHz	4	8.5 kHz	8 16	Stereo Mono

Note: CD-DA tracks will always follow CD-ROM XA tracks

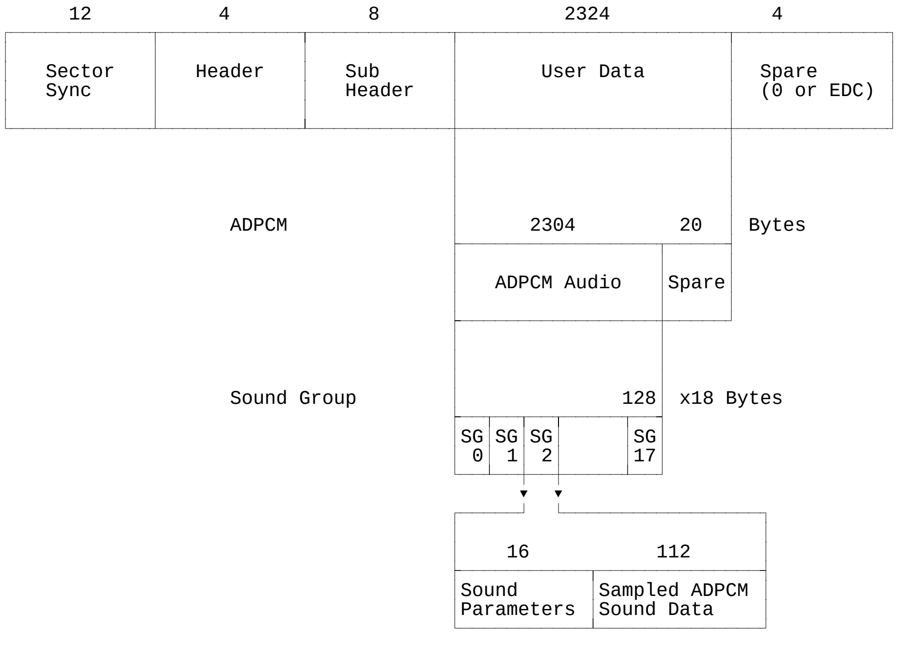
IV.3 **Audio Sector Data Format**

3.1 **General**

The data field of an audio sector (see II.4.6) contains three fields:

- a Subheader containing the File Number, Channel Number, Submode, and Coding Information bytes;
- an audio block data area containing 2304 bytes; and
- a padded field containing 20 bytes with value zero.

Figure IV.3 **Audio Sector data field**



3.2 Subheader Bytes

See II.4.3.2 for general definition of the Subheader bytes.

3.2.1 File Number

See II.4.3.2.1.

3.2.2 Channel Number

It should be noted that for audio sectors, the Channel Number can only have the value of 0 to 15.

The audio channel selection for the ADPCM decoder is controlled by a separate 16-bit audio channel selection register. Any audio channels selected by the 32-bit channel selection register but not the 16-bit register are transferred to the host system along with the other (non-audio) channels. The audio channel selected by the 16-bit register is transferred to the audio processor.

3.2.3 Submode

The submode byte has to be encoded as given in Figure IV.4 for an audio sector.

Figure IV.4 **Encoding for the Submode byte of an Audio sector**

bit number	7	6	5	4	3	2	1	0
bit value	x	x	1	x	0	1	0	x

where x is defined in II.4.3.2.3

It is recommended to set bit 6 (RT) to 1.

3.2.4 **Coding Information**

The Audio Coding Information byte is defined as follows:

Figure IV.5 **Audio Coding Information byte layout**

0	Emphasis	Bits per Sample	Sampling Frequency	Mono/Stereo
bit 7	6	5 4	3 2	1 0

Reserved
bit 7 = Set to zero

Emphasis (see IV.4.2)
bit 6
0 = Emphasis off
1 = Emphasis on

Number of Bits per Sample
bits 5-4
0 0 = 4 bits
0 1 = reserved
1 0 = reserved
1 1 = reserved

Sampling Frequency
bits 3-2
0 0 = 37.8 kHz
0 1 = 18.9 kHz
1 0 = reserved
1 1 = reserved

Mono/Stereo
bits 1-0
0 0 = Mono
0 1 = Stereo
1 0 = reserved
1 1 = reserved

3.3 **Audio Block**

The Audio Block field bytes B_{24} to B_{2327} of a CD-ROM XA audio sector consists of 2304 bytes. The Audio Block is further subdivided into 18 Sound Groups (SG_{00}...SG_{17}) of 128 bytes each. The Sound Groups have to be encoded in sequential order (see Figure IV.5).

Figure IV.6 **Ordering of Sound Groups**

Audio Block	Sound Group
B_{24} . . . B_{151}	SG_{00}
. . .	SG_{xx}
B_{2200} . . B_{2327}	SG_{17}

3.4 **Sound Group**

A Sound Group is divided into two parts:

- Sound Parameters (SP) : 16 bytes
- sampled audio data : 112 bytes

The bytes S_i in the Sound Group (SG) are indexed per SG from 0 to 127.

Figure IV.7 gives the layout of a Sound Group.

Figure IV.7 **Sound Group layout**

Sound Group byte numbers	meaning
S_0 . S_{15}	Sound Parameter bytes
S_{16} . . . . S_{127}	sampled audio data bytes

The Sound Parameter bytes each contain the two sound parameters, i.e., the Range (R) and Filter (F) (see IV.4.3).

The Sound Parameter (SP) has the value given by:

$$SP = 16 * F + R$$

where the Range (R) and Filter (F) positions in the Sound Parameter byte are shown in Figure IV.8.

Figure IV.8 **Sound Parameter byte layout**

Filter (F)	Range (R)
bits 7 to 4	bits 3 to 0

3.5 Level B and Level C Audio

A Sound Group (SG) consists of 8 Sound Units (SU^j) where j = 0 to 7. A Sound Unit consists of 2 identical Sound Parameter bytes of 8 bits and 28 Sound Data nibbles (SD^j_k) of 4 bits.

The Sound Parameters (SP^j) for the Sound Units (SU^j) are encoded in S_0 to S_{15} (see IV.3.4) and j is the Sound Unit index. Here,

$$S_i = SP^j$$

where the relation between i and j is given by Figure IV.8.

The 4 bit Sound Data (SD^j_k) from two Sound Units are combined into a Sound Byte (SB^l_k) where k is the sequential sampling order indexed from 0 to 27 in the following manner:

$$SB^l_k = SD^j_k + 16*SD^{j+1}_k \quad \text{(see Figure IV.9)}$$

where $j = 2 * l$ and
$l = 0$ to 3

These Sound Bytes (SB^l_k) are encoded in S_{16} to S_{127}. Here,

$$S_i = SB^l_k$$

where $i = 16 + l + 4*k$

In mono, the Sound Units are encoded sequentially i.e. SU^0, SU^1, SU^2...SU^7. In stereo, the left signal is given by SU^0, SU^2, SU^4, and SU^6 and the right signal is given by SU^1, SU^3, SU^5 and SU^7. The Sound Units are encoded in sequential pairs i.e. SU^n and SU^{n+1} are encoded together where n = 0, 2, 4 and 6.

Figure IV.9 shows the Sound Parameter layout for audio and Figure IV.10 gives the Sound Data layout for these audio levels.

Figure IV.9 **Sound Parameter Layout**

Sound Group byte number (= i)	Sound Unit number (= j)
0	0
1	1
2	2
3	3
4	0
5	1
6	2
7	3
8	4
9	5
10	6
11	7
12	4
13	5
14	6
15	7

Figure IV.10 **Sound Data Layout**

Sound Group byte number (= i)	Sound Unit number (= j)	Sound Sample number (= k)
16	1 and 0	0
17	3 and 2	0
18	5 and 4	0
19	7 and 6	0
20	1 and 0	1
21	3 and 2	1
.	.	.
.	.	.
126	5 and 4	27
127	7 and 6	27

3.6 **Audio Sector Interleaving**

Figure IV.11 **Audio Sector Interleaving**

	Relative Sector Number																
Level	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
B stereo	*				*				*				*				*
B mono	*								*								*
C stereo	*								*								*
C mono	*																*

Figure IV.11 shows possibilities for the audio sector interleaving for the case of a real-time bit being set to one. The interleave factors are fixed for each quality level and refer to the audio sectors with the same file number and channel number; the other audio sectors must have another file number or channel number.

The audio sector interleaving is not necessarily applicable if the real-time bit is zero for audio sectors.

IV.4 **Audio Data Encoding**

4.1 **General**

The principal feature of the ADPCM system (see Figure IV.12) is that it provides multiple prediction filters in order to respond effectively to fluctuations in the frequency distribution of the signal. This system employs near instantaneous compression to compress the dynamic range. The prediction filters and noise-shaping filters of the system switch simultaneously in order to minimize the energy of output noise.

Only first order and second-order digital filters are used in the encoder.

The encoder selects which predictor is most suitable in the manner defined below.

- The predictor adaptation unit in Figures IV.12 and IV.13 compares the peak value of the prediction errors sampled over 28 samples in a sound unit and then selects the filter which generates the minimum peak.

- Prediction errors which are chosen are gain-controlled (normalized by their maximum value) and noise shaping is executed at the same time.

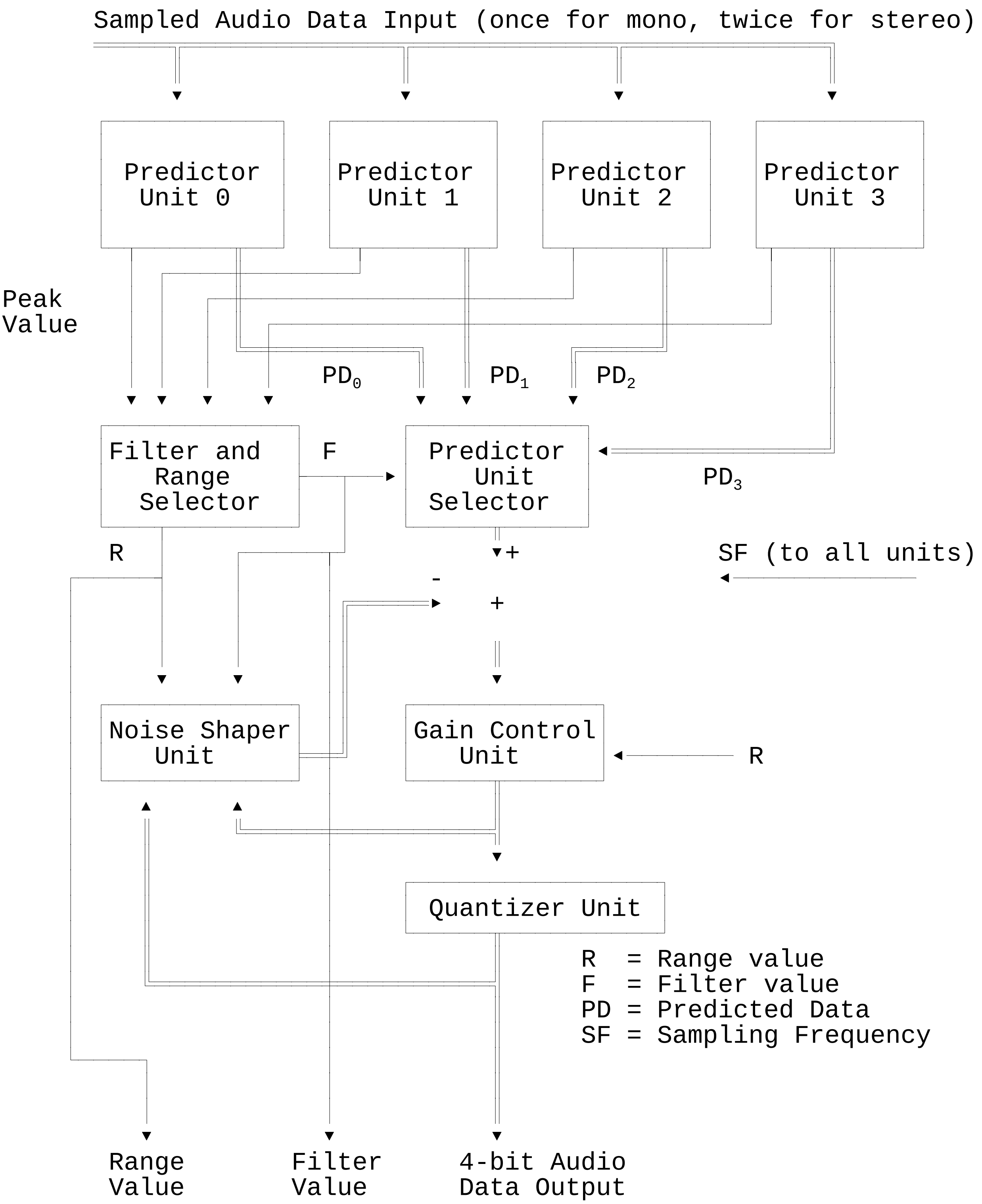
As a result of the above mentioned strategy, first-order and second order filters are used for signals in the low and middle frequencies and the straight pass through filter is used for high frequency signals. This is done in order to obtain a high instantaneous Signal/Noise ratio.

Different strategies of selecting filters are allowed. The above set is only one approach.

All data processing has to be done in 2's complement code with a sufficient number of bits to avoid extra quantization distortion.

The quantizer output of audio data is a rounded value.

Figure IV.12 **ADPCM Decoder block diagram**



4.2 **ADPCM Encoder**

Figure IV.12 gives the general configuration of the ADPCM encoder which needs to be done once for mono and twice for stereo.

If the input data has been recorded with emphasis (see the CD-DA specification, page 1A) then the emphasis bit in the coding information byte has a value of one; otherwise it has the value zero.

There are four predictor units indexed here as 0 to 3 (see Figures IV.12 and IV.13).

The filter and range selector generates the values called Filter (F) and Range (R) (see IV.4.3) depending on the peak values from the predictor units. The predictor unit with the lowest peak value is selected.

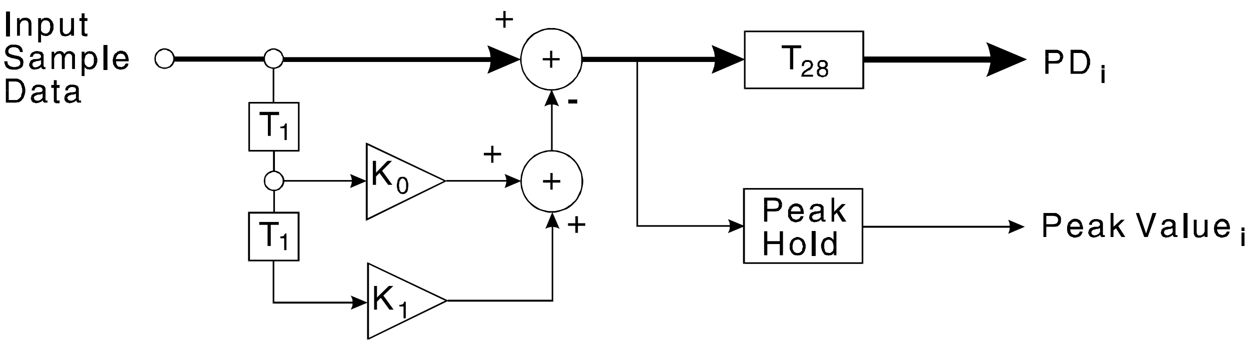
The predictor selector unit sends the chosen data to the gain control unit. The gain control unit multiplies the input data value with the gain factor 2^R where R is the Range value (see IV.4.3).

The purpose of the noise shaper unit (see Figure IV.14) is to compensate for extra noise introduced in the predictor unit.

The output value is rounded to 4 bits of audio data in the quantizer.

IV.4 Audio Data Encoding

Figure IV.13 **Predictor Unit**



T_1 = one sample delay

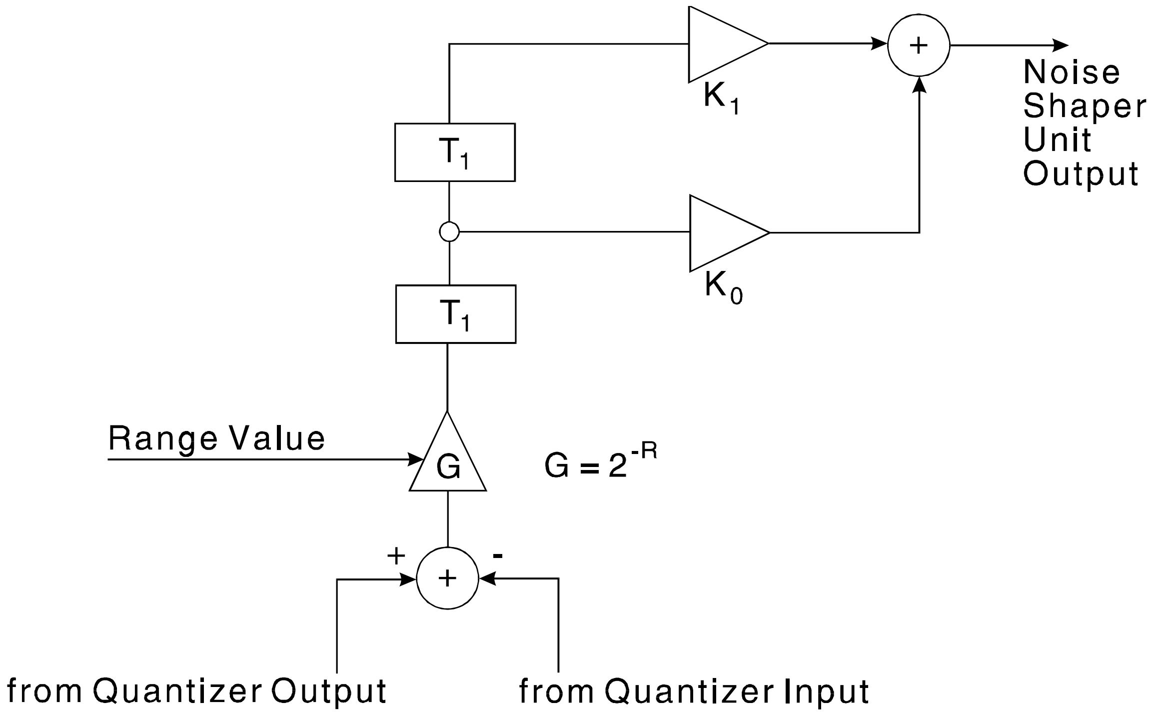
T_{28} = 28 sample delay

K_0, K_1 = gain factors and depend on the filter selected (see IV.4.3)

i = 0 to 3

IV.4 Audio Data Encoding

Figure IV.14 **Noise Shaper Unit**



T_1 = one sample delay

K_0, K_1 = gain factors and depend on the filter selected values.

4.3 Sound Parameters

Figure IV.15 shows the values of the gain for the various filters of the Sound Parameter. The range value has to be between 0 and 12.

Figure IV.15 **Gain Values for Sound Parameter Filters**

Filter	K_0	K_1
0	0 0.000000*	0 0.000000*
1	0.9375 0.111100*	0 0.000000*
2	1.796875 1.110011*	-0.8125 -0.110100*
3	1.53125 1.100010*	-0.859375 -0.110111*

(*) Base 2 rational numbers

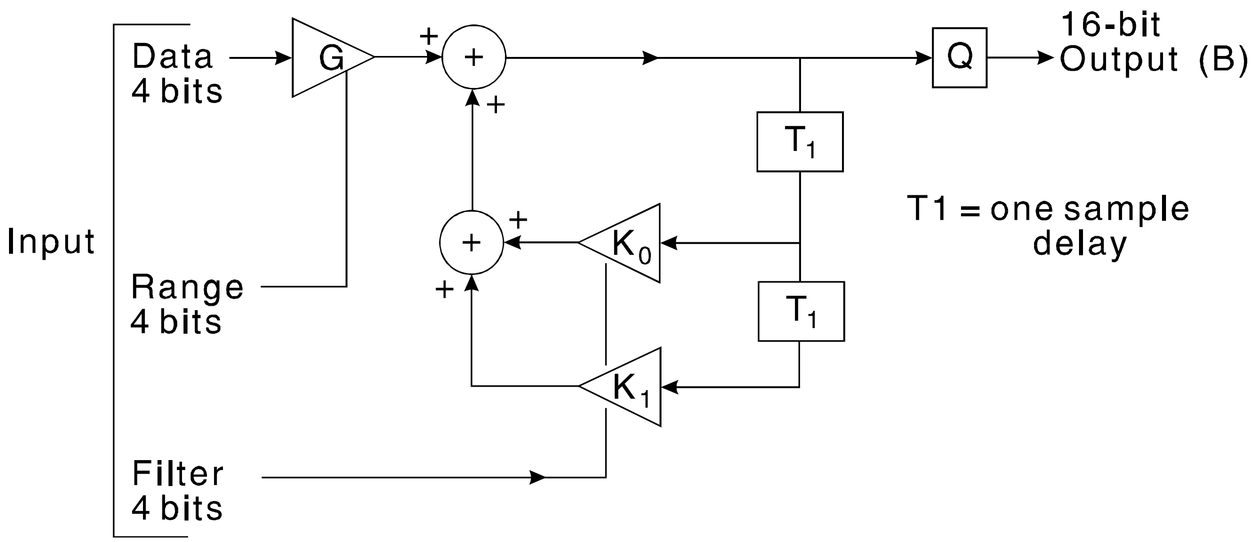
IV.5 **ADPCM Decoder**

The ADPCM decoder (see Figure IV.16) converts the CD-ROM XA audio sector encoded data to 2's complement 16-bit PCM coded audio (B). The sampling frequency depends however on the audio quality level used (see IV.2).

This unit has to buffer the CD-ROM XA audio data in order to decode the time-multiplexed signals.

A 4-bit audio data word is converted into a 16-bit word by the ADPCM decoder (see Figure IV.16).

Figure IV.16 **ADPCM Decoder**



The multiplier G adjusts the gain of the system by means of the range value so that the audio data is decoded into 16-bit words.

The multiplier algorithm is:

Word value = Audio data value * $2^{(12-R)}$

In each case the Range (R) is:

Range = Range data value (see Figure IV.14)

The filter data controls the coefficients K_0 and K_1 of the filter. These coefficients are the same as those used for the encoder (see Figure IV.13)

The output data is quantized to 16 bits by the quantizer Q. If overflow occurs then these must be limited (clipped) to the maximum values.

This page is intentionally left blank

Table of Contents

		Page
V.1	General Video Encoding and Display	1
V.2	Video Coding Information Byte	2
V.3	Video Data Sequence	3
V.4	Video Data Element format	4
V.5	CLUT data format	5

This page is intentionally left blank

V.1 **General Video Encoding and Display**

This chapter specifies the CD-ROM XA Extended Video Mode (EVM) data format as defined in Chapter II.4.3.2.4. The Color Look-up Table (CLUT) coding for CD-ROM XA is intended for CD-ROM XA systems which are able to display CLUT coded pictures.

It is recommended that a CD-ROM XA display system supports:

Resolution and/or	Horizontal ≥ 640 Horizontal ≥ 320	Vertical ≥ 480 pixel Vertical ≥ 200 pixel
CLUT	8, 4, 2 and 1	
RGB D/A	≥ 6 bit resolution per color	

A picture can be a full image of (640 x 480 or 320 x 200) or a rectangular part of the image. The CLUT coded picture is represented by picture pixel data recorded in CD-ROM XA Video sectors and by a CLUT table recorded in one CD-ROM XA Data sector (Form 1 only). The picture may be recorded together with other data in either Real Time sectors[1] or non-real time sectors.

The CLUT data is to be read and processed by application software. The picture pixel data also has to be processed by application software.

[1] Bit 6 in the Submode byte of the Subheader set to 1.

V.2 Video Coding Information byte[2]

The Coding Information byte in the Subheader for EVM Video coding is defined as follows:

	1	0	Resolution			Picture Coding		
Bit:	7	6	5	4	3	2	1	0

Resolution:

bits 5 4 3		
0 0 0	320 x 200	Hor x Ver
0 0 1	640 x 480	Hor x Ver
x x x	reserved	

Picture coding:

bits 2 1 0		
0 0 1	CLUT1	1 color per pixel
0 1 0	CLUT2	4 colors per pixel
0 1 1	CLUT4	16 colors per pixel
1 0 0	CLUT8	256 colors per pixel
x x x	reserved	

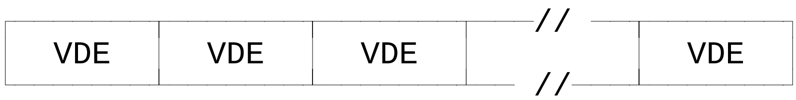
[2] Defined in the CD-I Full Functional Specification, chapter V.

V.3 Video Data Sequence

The Video Data Sequence (VDSQ) is the basic data unit for the recording of picture pixel data. A VDSQ contains a packed sequence of Video Data Elements (VDE) with no gaps between the VDE's. Each VDE represents one or more pixels and shall have the same CLUT format within one VDSQ.

A VDSQ must start at the beginning of a sector and contains an integral number of video sectors. The last sector may have to be padded with zero's.

Video Data Sequence:



Each VDE represents a CLUT index and is in most implementations identical to the pixel code in the display memory. All pixel data of a picture is stored in one single VDSQ.

The order of elements in a VDSQ respresenting a picture is ordered in rows from top to bottom of the picture and each row is ordered from left to right. Rows follow each other with no gaps or padding between them.

If the picture is a partial image then it is the responsibility of the application to know the number of pixels per row and number of rows for the picture and to load these data into the correct positions of the display memory.

V.4 Video Data Element format

The VDE format for the different CLUT coding methods is defined below. Ci represents the index of the CLUT element for pixel number = i in a VDSQ, starting from zero.

Pixel data format of the VDE for CLUT8:

	Ci
bit	7 … 0

Pixel data format of the VDE for CLUT4:

	Ci	Ci+1
bit	7 … 4	3 … 0

where i mod 2 equals 0.

Pixel data format of the VDE for CLUT 2:

	Ci		Ci+1		Ci+2		Ci+3	
bit	7	6	5	4	3	2	1	0

where i mod 4 equals 0

Pixel data format of the VDE for CLUT1:

	Ci	Ci+1	Ci+2	Ci+3	Ci+4	Ci+5	Ci+6	Ci+7
bit	7	6	5	4	3	2	1	0

where i mod 8 equals 0.

V.5 CLUT data format

All data of the CLUT is recorded into one CD-ROM XA data (Form 1 RT or non-RT) sector.

The CD-ROM XA data sector format is only specified for CLUT data which defines the color values of the RGB components for each CLUT element.

A CLUT data sector contains packed sequences of one or more CLUT Data Sequences (CDSQ), which consist of a Header part (4 bytes) and a part containing the CLUT Data Elements (CDE). There may be no gap or padding bytes between the CDSQ's and they all have to fit into one sector of 2048 bytes.

The CLUT is divided into banks of 64 elements, with index addresses #0..#63, #64..#127, #128..#191 or #192..#255 to the data elements. CLUT1, CLUT2 and CLUT4 uses only the first bank #0..#63, but CLUT8 uses all banks.

A CDSQ can contain up to 64 continuous sequences of CLUT elements, but may only contain elements from one and the same bank. A CDSQ for #128 to #159 may follow a CDSQ for #0 to #15, skipping #16 to #127.

Example of a Data Sector layout

CDSQ #0..#15	
CDSQ #128..#159	
= =	
CDSQ #192..#193	
Padding with zero	(up to 2048 bytes)

CLUT table Data Sequence (CDSQ):

Header	CDE	CDE	CDE	//	CDE

V.5 CLUT data format

CDSQ Header format (4 bytes)

Byte Position		
1	First element in CLUT	(0..255)
2	Last element in CLUT	(0..255)
3	Number of elements in the CDSQ	(1..64)
4	bit 7 : End of CDSQ 0: More CDSQ's follow 1: Last CDSQ in sector	
	bit 6-2 : reserved (= zero)	
	bit 1-0 : CLUT bank number (0..3).	

CLUT Data Element (CDE) format (4 bytes)

Byte Position			
1	INDEX	: Index of element in bank +128	(128..191)[3]
2	RED	: Intensity level of red	(0..255)
3	GREEN	: Intensity level of green	(0..255)
4	BLUE	: Intensity level of blue	(0..255)

The Red, Green and Blue components should be encoded such that the black level is at R,G,B, = 0,0,0 and the peak white level is at R,G,B = 255, 255, 255.

[3] Conforms to CD-I Full Functional Specification, chapter V.

1. **Introduction**

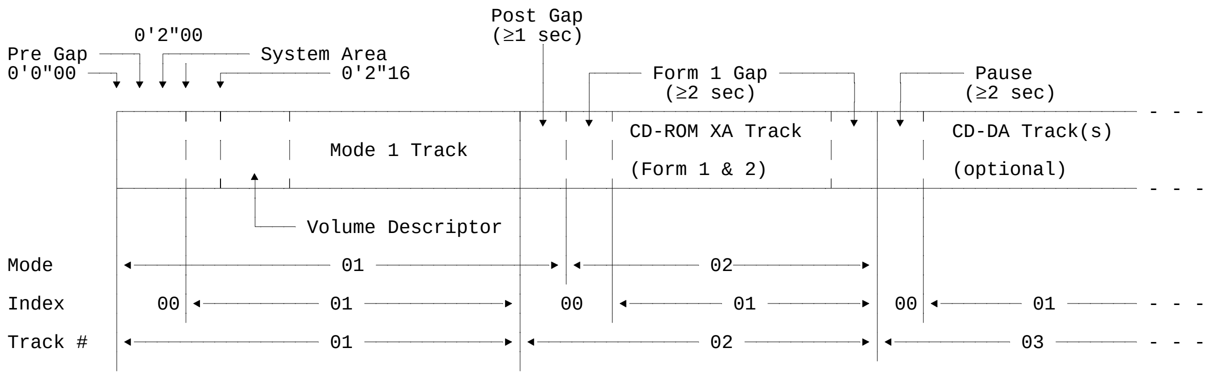
Some early CD-ROM drives are not able to read discs with data tracks recorded in Mode 2.

If it is desired that discs be read by CD-ROM drives capable of reading Mode 1 tracks only, a special CD-ROM XA disc can be created by adding a CD-ROM Mode 1 track as a first track. This optional format is described in this appendix.

For compatibility reasons, some restrictions have to apply.

2. **Disc Format**

Figure AI.1 **Track layout of Disc with Mode 1 Track**



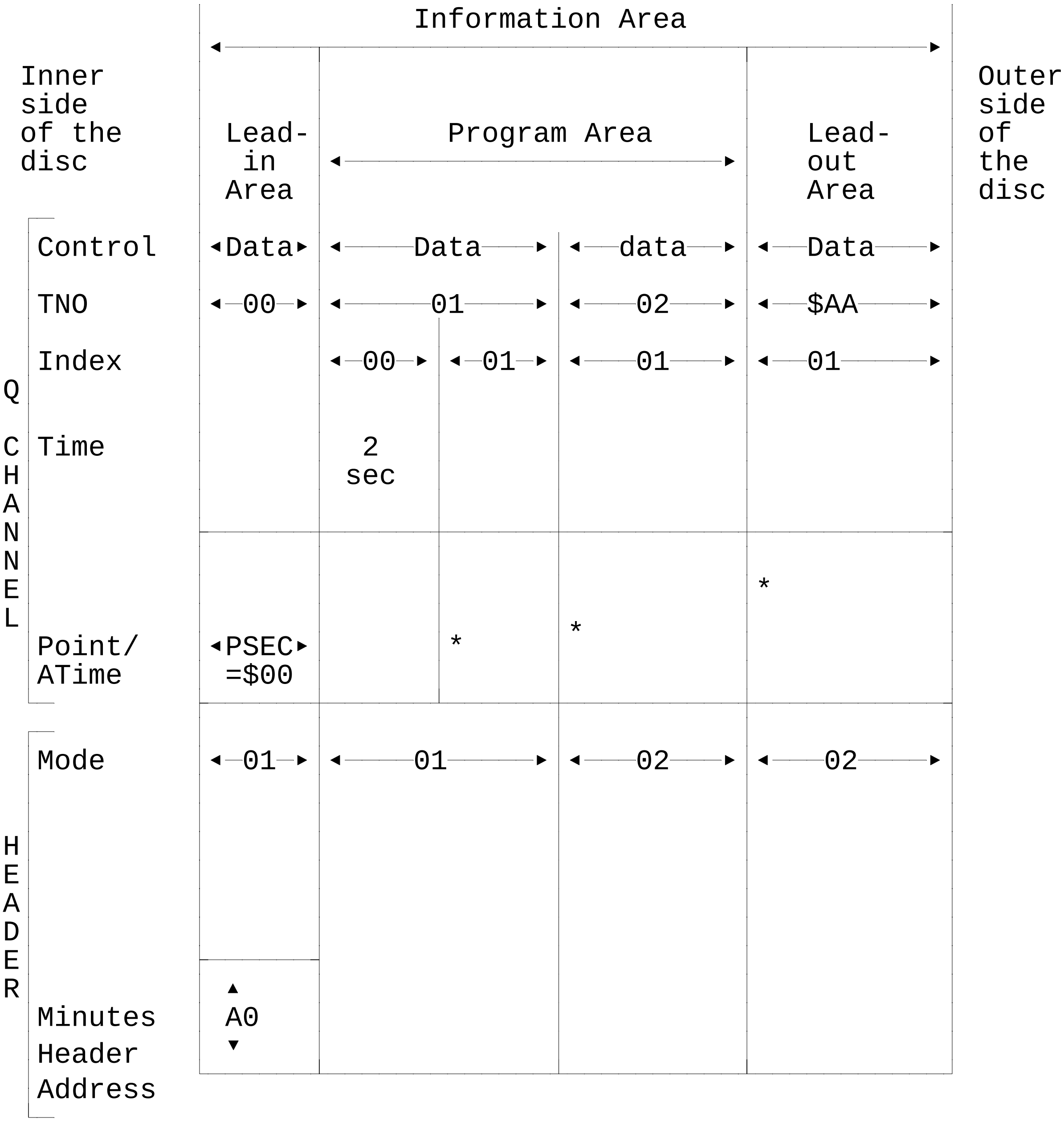
This type of disc is identified by the subcode Q channel in the lead-in area with POINT = $A0, PSEC = #00 amd PFRAME =$00.

The disc consists of CD-ROM Mode 1 and CD-ROM XA Mode 2 tracks (recommended one track each), and CD-DA tracks if desired. The CD-ROM XA track(s) shall follow the CD-ROM Mode 1 track.

This track layout is to be used when the disc contains data which must also be read by CD-ROM drives which can only read Mode 1 tracks.

AI.2 Disc Format

Figure AI.2 **Example of the general structure and coding of the TOC for a disc with one Mode 1 and one CD-ROM XA track**



* These points as specified in the Q-channel give the start position of the first and the last track or lead-out area in ATime with an accuracy of ± one second.

3. **Data Retrieval Structure**

There are three kinds of CD-ROM discs:

(1) - CD-ROM discs with only Mode 1 tracks
(2) - CD-ROM XA discs with only CD-ROM XA tracks (Mode 2 Form 1 or Form 2)
(3) - CD-ROM discs with both Mode 1 and CD-ROM XA tracks (described in this Appendix)

Note: CD-DA track(s) can follow data tracks, if desired.

Discs (1) recorded with only Mode 1 tracks correspond to existing CD-ROM discs and are already described by the CD-ROM and ISO 9660 Specifications.

Discs (2) recorded with CD-ROM XA tracks are described in this System Description of CD-ROM XA. The file structure is based upon the file structure for CD-ROM discs (1) in Mode 1 specified in ISO 9660 with extensions listed in chapter III.

Note: Mode 1 and Mode 2 Form 1 sectors are logically equivalent, with both 2048 bytes per sector.

For discs (3) with both Mode 1 and CD-ROM XA tracks, compatibility with CD-ROM Mode 1 only drives is a concern. It is possible within the existing ISO 9660 specification to arrange the file structure layout so that Volume Descriptors, Path Tables, Directories and some files are in Track 1 with Mode 1 sectors, whereas the CD-ROM XA files are in Track 2 with Mode 2 Form 1 or Form 2 sectors. Although the file structure is the same, files written in Mode 2 Form 1 or Form 2 tracks are unreadable in drives that cannot read Mode 2 Form 1 and Form 2 sectors.

By using discs with both Mode 1 and CD-ROM XA tracks existing CD-ROM discs can be enhanced with CD-ROM XA capabilities without sacrificing compatibility with existing drives.

ADPCM audio files shall be recorded in the CD-ROM XA track only. In the case of a CD-ROM XA Disc with an optional Mode 1 track, character or textual data written in the CD-ROM XA track may be written in part or in full in the Mode 1 track. Graphic or video data written in the CD-ROM XA track may also be written in the Mode 1 track.

In the case of a CD-ROM XA Disc with an optional Mode 1 track, character or textual data written in the CD-ROM XA track may be written in part or in full in the Mode 1 track. Graphic or Video data written in the CD-ROM XA track may also be written in the Mode 1 track.

AI.3 Data Retrieval Structure

Any character, graphic or video data written in Mode 1 must also be written in CD-ROM XA tracks.

The Directory Record for the Start-up Directory however should be in the Mode 1 track, though the start-up files can be in the Mode 2 track.

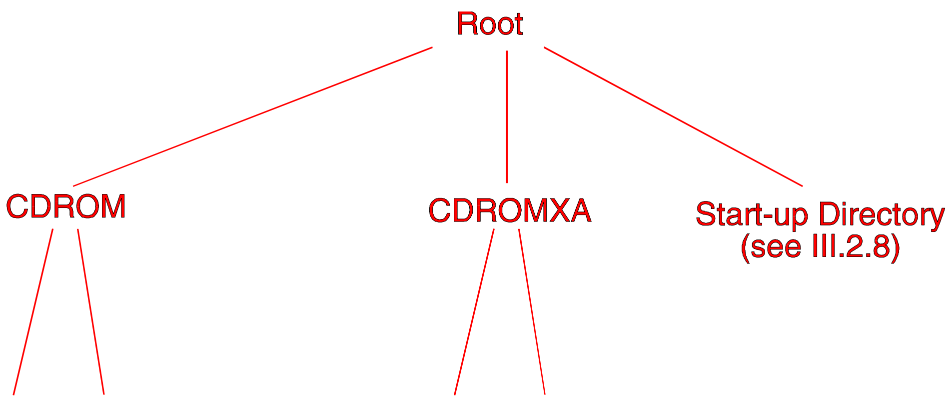
Files shall not cross the boundary between tracks with different modes.

Receiving systems are free to make hidden any files whose attributes indicate that they are recorded in Mode 2 when the discs are placed in drives capable of reading Mode 2.

As there are receiving systems that are not aware of the CD-ROM XA attributes in the system use field, it is recommended in this case to set the Existence bit of CD-ROM XA files to 1 (ref. ISO 9660 section 9.1.6).

If hiding of Mode 2 files by the receiving system is not implemented, then reading of these files may cause the application to fail.

Because of this, it is recommended to use a directory hierarchy with at first two subdirectories, **CDROM** and **CDROMXA**, for the Mode 1 track and the Mode 2 track respectively. The **Root** directory must be in the Mode 1 track as well. In this way it is easy to identify how files are recorded.



Listing of registered Directory Identifiers

For current and new registrations, please contact the address given in the Preface (page i).